Contributions to Theoretical Economics

Volume 4, *Issue* 1 2004 *Article* 6

Players With Limited Memory

Steffen Huck* Rajiv Sarin†

*University College London, s.huck@ucl.ac.uk

†Texas A&M, rsarin@econmail.tamu.edu

Players With Limited Memory*

Steffen Huck and Rajiv Sarin

Abstract

This paper studies a model of memory. The model takes into account that memory capacity is limited and imperfect. We study how agents with such memory limitations, who have very little information about their choice environment, play games. We introduce the notion of a Limited Memory Equilibrium (LME) and show that play converges to an LME in every generic normal form game. Our characterization of the set of LME suggests that players with limited memory do (weakly) better in games than in decision problems. We also show that agents can do quite well even with severely limited memory, although severe limitations tend to make them behave cautiously.

KEYWORDS: bounded rationality, games, memory, maxmin

*We wish to thank two anonymous referees for very helpful comments. The first author acknowledges financial support from the Economic and Social Research Council (UK) via the Centre for Economic Learning and Social Evolution. email: Steffen Huck, s.huck@ucl.ac.uk; Rajiv Sarin, rsarin@econmail.tamu.edu.

1 Introduction

It is widely acknowledged that memory limitations affect behavior. In this paper we develop a model of memory which explicitly takes into account that memory capacity is limited. Even this limited capacity is imperfect, in the sense that arbitrary items in memory may be forgotten. We study how individuals with such memory limitations play games.

We suppose that a player's memory contains, for each of her strategies, a record of a finite number of the most recent payoffs obtained. New information leads to the deletion of old information. Players do not actively select which payoffs to store in their memories. Nor do they store "processed" information in the form of summary statistics of their past experiences. The players are not sophisticated. By not storing processed information, they are probably not making efficient use of their limited memory. When they think back they only recall the most recent payoffs they have experienced from each strategy.

The players repeatedly play the same normal form game in which each knows only the strategies available to her. At each stage the players first choose a strategy and then receive a payoff. The payoff each player obtains depends on the chosen strategy profile. At the time of making their choices the players do not know the choices of others. The players are not assumed to know the payoff functions of other players, or even the strategies available to their opponents. They need not even observe the strategy profile chosen by their opponents. In fact, they do not even need to know that they are playing a game.

The players choose among their strategies on the basis of what they remember about the performance of different strategies. We suppose that how an agent evaluates any particular strategy is monotonic in payoffs. Roughly speaking, monotonicity requires that an agent evaluates a strategy as better if it has given higher payoffs. An example of a monotonic rule is the evaluation of each strategy according to the average payoff that the strategy has received in the past. Other examples of monotonic rules are evaluations according to the minimum payoff or the maximum payoff that the strategy has received in the remembered past. We suppose that the players choose, at each time, the strategy they evaluate as being the best. That is, we assume the players are myopic.[1]

Our first result concerns the model in which agents have limited memory

[1] Myopia may also explain why the players do not attempt to store summary statistics: The efficient storage of information, for possible use in the future, cannot possibly be a concern of a myopic agent.

capacity but are not forgetful in the sense that they do not forget arbitrary items in their memory.[2] We define a *limited memory equilibrium* (LME) as a strategy profile which is associated with an absorbing state of the dynamics describing how memory and play evolve over time. We show that a strategy profile is an LME if and only if each player obtains at least her maxmin payoffs.[3] We are also able to show that play must converge to an LME, starting from any initial state, if players use any monotonic evaluation rule. This result contrasts with that obtained by Sarin (2000) who shows that, in a decision problem, a player converges to choose her maxmin strategy.[4] Intuitively, the (weakly) superior performance of agents in games as opposed to decision problems arises for two reasons. On the one hand, payoffs are deterministic in the games that we consider and on the other hand players may, unsuspectingly, influence and improve the nonstationary environment they face, whereas in the multi-armed bandits they cannot possibly alter the stationary environment with which they are confronted.

The remainder of the paper is concerned with agents who have a limited memory capacity and who are forgetful. A forgotten payoff is replaced by a possible payoff. We refer to a strategy profile that forgetful players may play as a *stable limited memory equilibrium* (SLME). The set of strategy profiles that are SLME is contained in the set of strategy profiles that are LME. In general, the set of SLME depend on the specific monotonic evaluation rule used by the players. In particular, we show that the set of SLME depend on the cardinal properties of payoffs. We proceed by providing results on specific monotone rules whose behavior depends only on the ordinal properties of the remembered payoffs. We show that if players use the maximum rule, and have a large enough memory, then the unique SLME in games of common interest is the strategy profile that induces the Pareto–optimal outcome. If players use the minimum rule we show that the strategy profile in which each player plays her maxmin strategy is an SLME and that it is the unique SLME if maxmin play constitutes a Nash equilibrium. This implies, for example, that such players will choose the risk–dominant equilibrium in 2x2 games in which risk- and payoff–dominance conflict.

Next we introduce a class of games, defined by *iterated uniform dominance* (IUD), in which SLME arise independently of the particular monotone

[2] This model is studied in decision problems in Sarin (2000).

[3] The maxmin payoff for a player is the highest minimum payoff she can guarantee herself when using only pure strategies. The strategy that ensures a player her maxmin payoff is referred to as her maxmin strategy.

[4] For other justifications of maxmin strategies, see Barbera and Jackson (1988), Gilboa and Schmeidler (1989) and Sarin and Vahid (1999).

evaluation rule used. We say a strategy s is uniformly dominated if there exists another strategy whose minimum payoff is larger than the maximum payoff that may be obtained from s. Eliminating a uniformly dominated strategy may make another strategy uniformly dominated. If we iteratively eliminate all of the strategies that are uniformly dominated then we obtain the set of strategies that survive the process of IUD. If this set is a singleton, we say that the game is solvable by IUD.[5] For the class of uniformly dominance solvable games, we find that the unique Nash equilibrium is an SLME regardless of evaluation rules and memory capacities. For the larger class of dominance solvable games, we show that if players use the minimum evaluation rule, the Nash equilibrium is an SLME.

In contrast with this paper and Sarin (2000), Sela and Herreiner (1999) consider the case where the agent remembers only what has happened in a finite number of the most recent periods. The different models of memory are coupled with different assumptions regarding what the agent knows and what she observes. In Sarin the agent knows only her available strategies and observes only the payoff from the chosen strategy. In Sela and Herreiner the agent knows the payoff matrix including the possible states of the world and observes the chosen state of the world. In the terminology of the current literature on learning, the Sarin model can be considered a "reinforcement" model of limited memory whereas the Sela and Herreiner model can be considered a "belief" model of limited memory.

A richer "belief" model of memory is considered by Mullainathan (1998) who focuses on rehearsal (recalling a memory increases future recall probabilities) and association (events more similar to current events are easier to recall). The agents are assumed to know more about their environment than in Sela and Herreier and use Bayes' rule to update their beliefs. Mullainathan uses his model to explain certain regularities in income and consumption data, and some aspects of asset pricing. Papers that have implicit belief models of memory include Hurkens (1995) and Young (1993) who assume that players are selected from populations to play a game. When called upon to play, a player samples a fixed number of the choices her opponents made in the recent past. This finite sample forms the player's memory of the past play of the game. The agent best replies to her memory. After making a choice the individual forgets everything. The model is used to explain the

[5]This solution concept has been independently suggested by Friedman and Shenker (1998), who refer to the set of strategies surviving IUD as the serially unoverwhelmed set. Chen (1999) has found the solution concept to be of use for explaining experimental data of public goods pricing mechanisms and Greenwald, Friedman and Shenker (1998) have studied its relevance in network contexts.

evolution of mutually consistent behavior in a population. More sophisticated players with memory limitations have been studied by Aumann and Sorin (1989), Lehrer (1988, 1994) and Sabourian (1998).

Osborne and Rubinstein (1998) consider "reinforcement" type model of memory in which agents have about the same level of sophistication as the agents we consider. They know the set of available actions but do not know that they are playing a game. They choose each action a fixed, finite number of times and evaluate each action according to the sum of payoffs it has given. The action thought to be the best is chosen. They introduce an equilibrium notion relevant for such players and study its properties. To compare our model with theirs, it is useful to consider the initial memory of our players. One possibility is that their initial memory arises exactly according to the Osborne and Rubinstein procedure, in which each agent chooses each action a fixed, finite number of times. In contrast to Osborne and Rubinstein, we allow that agents evaluate the payoffs in their memory in a large variety of ways. The evaluation rule they consider is a specific monotone rule. Also, in contrast to their static equilibrium notion, the equilibrium notion(s) introduced in this paper are derived as the limiting (absorbing) states of the large class of dynamics we consider and in which the memories of the players are endogenously evolving. Osborne and Rubinstein do not provide an analysis of how the equilibrium comes about. This is done in a paper by Sethi (2000) who describes the dynamics in a large population setting.

This paper is organized as follows. The next section presents the basic model. Section 3 characterizes LME. Section 4 introduces additional forgetfulness and focuses on SLME, and Section 5 provides results about SLME in games solvable by IUD and by iterated strict dominance. Section 6 discusses possible extensions and limitations and Section 7 concludes.

2 The Model

Consider a finite normal form game $\Gamma = (I, S, u)$, where $I = (1, ..., n)$ denotes the set of players with typical element i. $S = \times_i S^i$ is the set of possible strategy profiles in the game and S^i is player i's set of strategies. A typical element of S is given by $s = \left(s^i, s^{-i}\right)$ where $s^i \in S^i$ denotes the strategy of player i and s^{-i} specifies the strategies of players other than player i. S^{-i} is the set of strategy combinations available to players other than i. We shall suppose that player i has J_i available strategies and sometimes we shall number the strategies from $s^i_1, ..s^i_{J_i}$. By u we denote the payoff players receive from alternative strategy profiles. Specifically, the payoff that player i obtains in the strategy profile s is given by $u^i(s) = u^i(s^i, s^{-i})$. That is, $u^i : S \to \Re$, and $u = \times_i u^i$. Clearly, $u : S \to \Re^n$.

The players have limited memories. Each player i associates with each strategy H payoffs. These will be the most recent H payoffs the player has obtained from the choice of the strategy. Let $m\left(s_j^i, t\right)$ denote the H payoffs that player i associates with strategy s_j^i at time $t = 0, 1, 2, \ldots$ which when no confusion may result we shall simplify as m_j^i. Let $m_h(s_j^i, t)$ be the hth element of this vector, $h = 1, 2\ldots, H$. Furthermore, let $m^i(t) = \left(m\left(s_1^i, t\right), \ldots, m\left(s_{J_i}^i, t\right)\right)$ be the state of player i's memory at time t. We shall suppose that the initial contents of the memory of the players is given and satisfies the condition that for each player i, each $s^i \in S^i$, and each h, $m_h(s^i, 0) = u^i(s^i, s^{-i})$ for some s^{-i}. That is, each payoff that the agent has in her initial memory of strategy s^i is a payoff that s^i can actually obtain. This restriction on initial memory can be thought of requiring a certain degree of realism, and it can be justified by assuming that each agent has been endowed with a number of random observations from the actual game matrix. Alternatively, we may suppose that the initial conditions are realized as a consequence of a period in which the players experiment with all of their strategies.

Such memory allocation supposes that players recall only their payoff experiences with different strategies, and that they recall only the most recent payoff experiences with any particular strategy.[6] Payoffs obtained earlier are forgotten. As some strategies may not have been chosen for a long time, such memory use implies that the decision maker recalls as many payoffs from recently chosen strategies as from those she has chosen only in the distant past.[7] Modelling memory allocation in this manner makes comparisons between different strategies straightforward: The decision maker has only to compare payoff vectors with the same number of elements. Considerations involved in comparing strategies regarding which the agent has different amounts of information do not have to be addressed in this model.

Each player evaluates her strategies according to a monotonic evaluation rule. An evaluation rule is monotonic if it evaluates a strategy as better if it has yielded higher payoffs in the (remembered) past.

[6] Observe that we assume that players experience no problems in retrieving objects in their memory.

[7] In Section 6 we discuss models of memory allocation in which the decision maker remembers a different number of payoffs from different strategies.

Definition 1 *An evaluation rule* $f : \Re^H \to \Re$ *is* monotonic *if whenever* $f(x) = f(y)$, *x not necessarily equal to y, then for all* $q \geq 0$, $q \in \Re^H$, $f(x+q) \geq f(y)$, *and if* $q > 0$ *then* $f(x+q) > f(y)$.

An example of a monotonic evaluation rule arises when a player evaluates a strategy according to the average payoff it has given her in her remembered past. We call this evaluation rule the average rule. Another monotonic evaluation rule arises when an agent evaluates each strategy according the minimum payoff it has given her in the remembered past. We shall call this the minimum rule. Yet another monotonic evaluation rule lets an agent evaluate a strategy according to the maximum payoff it has given her in the remembered past. We shall call this the maximum rule.[8]

We shall assume that at each time each player chooses the strategy which she evaluates as being best. That is, agents are assumed to be myopic. In the context of our model, myopia can be justified as the agents have very little information about the payoff functions they are facing and, as a consequence, experience large amounts of subjective uncertainty.[9] We shall suppose that if a player evaluates more than one strategy as being best, then the player will choose each of these strategies with positive probability bounded away from zero.[10] Let

$$B^i(m^i(t)) = \left\{ s^i_j : f(m(s^i_j, t)) = \max \left\{ f\left(m(s^i_1, t)\right), ..., f\left(m(s^i_{J_i}, t)\right) \right\} \right\}$$

be the set of strategies whose evaluation is the highest at time t. With this notation we can easily define a player's decision rule: At time $t+1$ player i plays strategy s^i_j with probability 0 if $s^i_j \notin B^i(m^i(t))$ and with some probability $p^i_j(t) > 0$ if $s^i_j \in B^i(m^i(t))$ where $\sum_j p^i_j(t) = 1$.

A state at time t is described by the contents of all players' memories at that time, i.e., by $m(t) = \times_{i \in I} m^i(t)$. Let $\mathcal{M}$ denote all possible constellations of m. Note that, $m(t)$ does not reveal which strategy profile will be played in period t. Rather it induces a distribution over the set of strategy profiles according to which players will play in t. The support of this distribution is given by $\mathcal{B}(m(t)) = \times_i B^i(m(t))$. This, in turn, induces a probability distribution over $\mathcal{M}$. The definitions of the game, of players'

[8] Pessimistic players (i.e. players who always "expect the worst") would tend to adopt the minimum rule, while optimistic players (i.e. players who always "expect the best") would tend to adopt the maximum rule.

[9] Sonsino (1998) has shown that "strong" uncertainty may lead a non–myopic agent to behave in a myopic manner. Ellison (1997) has studied situations in which a rational non–myopic player may behave in a myopic manner.

[10] Players are assumed to randomize independently of other players in such situations.

memories, evaluation and decision rules together define a Markov process P on a finite state space $\mathcal{M}$.

3 Limited Memory Equilibrium

The equilibrium concept we develop in this section is appropriate to describe the strategy profiles that players with limited memories will converge to play if each uses some monotonic evaluation rule.[11] We begin by characterizing the *absorbing states* of $\mathcal{M}$. Let $U^i(s)$ denote an H–vector consisting of $u^i(s)$ in every component. We say a game is generic if all payoffs any specific player can get are distinct.

Lemma 1 *In generic games a state m is absorbing if and only if*
(i) $\mathcal{B}(m)$ is a singleton, i.e., $\mathcal{B}(m) = \{s\}$ for some s, and
(ii) $m(s^i) = U^i(s)$ for all player i.

Proof *If*: From (ii) it follows that if players play s the state of memory does not change and (i) ensures that players do play s.

Only if: Suppose there was an absorbing state not fulfilling (i) or (ii). If (i) was not fulfilled there would be a positive probability for at least one player to experience different payoffs from one and the same strategy which, in generic games, implies a positive probability for her memory to change. If (ii) was not fulfilled but (i) was, $m(s^i)$ would change with probability 1 for some player i. $\square$

We shall refer to a strategy profile as a *limited–memory equilibrium (LME)* if it is played in an absorbing state.

Definition 2 *A strategy profile s is an LME if there exists an absorbing state m with $\mathcal{B}(m) = \{s\}$.*

LME need not exist in all games. For example, an LME fails to exist in the usual, non–generic, version of *matching pennies*.[12] It is, however, possible to see that LME exist in all generic games.[13] An easy way to see this is

[11] Different players may use different evaluation rules.

[12] If agents play one of the pure-strategy profiles for H periods, one of the agents (the loser) will evaluate her alternative strategy as at least as good as the currently used one such that $\mathcal{B}(m)$ will no longer be a singleton.

[13] A definition of equilibrium that would allow for existence in all games would require us to introduce a setwise analogue of the definition of LME, based on absorbing sets rather than absorbing states. A setwise equilibrium notion, however, would lead to significantly more cumbersome notation, without adding significantly to the insight. We choose, rather, to focus mainly on generic games where all absorbing sets are singletons as shown below.

to note that in generic games the maxmin strategy for each player is unique. The strategy profile in which each agent plays her maxmin strategy is an LME as the next result reveals. The result characterizes the set of LME in terms of payoffs the players obtain in them. Let $u^i_{\min}(s^i)$ denote the minimal payoff s^i can give to player i, i.e., $u^i_{\min}(s^i) = \min_{s^{-i}\in S^{-i}} u^i(s^i, s^{-i})$ and let $U^i_{\min}(s^i)$ be the H–vector consisting of $u^i_{\min}(s^i)$ in every component. The maximal payoff s^i can give is defined analogously and denoted by $u^i_{\max}(s^i)$. By $u^i_{\max\min} = \max_{s^i\in S^i} u^i_{\min}(s^i)$ we denote player i's maxmin payoff. The strategy yielding this payoff is denoted by $s^i_{\max\min}$ (as we focus in the following on generic games, we know that there is only one such strategy).

Proposition 1 *In generic games, a strategy profile s is an LME if and only if $u^i(s) \geq u^i_{\max\min}$ for all i.*

Proof *If*: Let $m(s^i) = U^i(s)$ and let $m(\widetilde{s}^i) = U^i_{\min}(\widetilde{s}^i)$ for all strategies $\widetilde{s}^i$ other than s^i and for all i. Note that $u^i_{\max\min} > u^i_{\min}(\widetilde{s}^i)$ as no two payoffs are equal. As $u^i(s) \geq u^i_{\max\min}$ it follows that $m(s^i) > m(\widetilde{s}^i)$ for all $\widetilde{s}^i$ and all i. By monotonicity of f it follows that $B^i(m) = \left\{s^i\right\}$ for all i. Hence, by Lemma 1 the result follows.

Only if: Suppose the opposite. That is, suppose that some player i gets less than her maxmin payoff in a strategy profile s which is an LME. This implies that $m(s^i) = U^i(s)$. Since $U^i(s) < m(s^i_{\max\min})$ it follows by monotonicity of f that $s^i \notin B^i(m)$. Hence, by Lemma 1 s cannot have been an LME.$\square$

Proposition 1 characterizes LME in terms of payoffs each player must obtain. This makes it easy to check whether a strategy profile is an LME. The result shows that the players cannot do "too badly" in any LME. Specifically, in an LME each player obtains a payoff at least as high as her maxmin payoff.

An example which illustrates Proposition 1 and which also reveals that LME do not have to be Nash equilibria, is obtained by considering the Prisoner's Dilemma game. The payoff from mutual cooperation is greater for each player than the maxmin payoff and hence mutual cooperation is an LME. For example, cooperation can be sustained in an LME when all players remember from the strategy "defect" only the mutual defection payoff.

In non–generic games, a pure Nash equilibrium is not necessarily an LME. To see this, simply consider a degenerate 2x2 game in which the row player receives the same payoff for all strategy combinations. The column player prefers *left* over *right* when the row player plays *up* and vice versa for *down*. This game has two Nash equilibria in pure strategies but neither is an LME. To see this consider the equilibrium $(up, left)$. Obviously, the row player's memory state is time–invariant. As the row player always assigns positive probability to both his strategies, there is always a positive probability that the column player's memory for strategy *left* will change. Hence, the equilibrium is not an LME.

While not every Nash equilibrium is necessarily an LME, every strict equilibrium is an LME. To see this, consider $u^i(s) > u^i(\widetilde{s}^i, s^{-i})$ for all $\widetilde{s}^i$ and for all player i. Now consider a state where $m(s^i) = U^i(s)$ for all i and where $m(\widetilde{s}^i) = U^i(\widetilde{s}^i, s^{-i})$ for all $\widetilde{s}^i$ and for all i. By monotonicity of f it follows that $\mathcal{B}(m) = \{s\}$ such that m cannot be left. Hence, the strict Nash equilibrium is also an LME.

In the next proposition show that players who use monotonic evaluation rules converge to some LME. The intuition of the result is simple. First we show that play cannot converge to an outcome that is not an LME (which is easy). In the next step we rule out cycles. The intuition for why we can do this is that agents switch from one strategy to another only when the evaluation of the currently used strategy has deteriorated. So a cycle would imply endless deterioration of the evaluations of each strategy that is used in the cycle. Given the finite nature of agents' memory this is clearly impossible. While this is straightforward and easy to show the proof is slightly more complicated since we also have to rule out cycles that results from randomizations where agents have the same evaluation for two or more strategies.

Proposition 2 *In generic games, starting from any initial state, play converges to an LME with probability 1.*

Proof First, we show that play cannot converge to any state that is not an LME. Suppose play converges to a state in which (at least) one player i gets a payoff below her maxmin payoff $u^i_{\max\min}$. Then her memory would contain only this payoff after at most H periods. But then the agent would evaluate her maxmin strategy as better, because she uses a monotonic evaluation rule. Hence play cannot converge to any state in which any player gets a payoff below her maxmin payoff.

Next, we argue that play cannot cycle among strategies. Suppose that play cycles between some strategies which includes strategy s^i_j for player i. Consider, first, cycles in which $f^i\left(m^i_j\right) \neq f^i\left(m^i_k\right)$ for all $j \neq k$ at any time. In this case, each time the agent returns to choose a strategy, the evaluation of it must have strictly declined. This is because a player only leaves a strategy after its evaluation has strictly declined, given that the evaluations of unplayed strategies stay unchanged. Hence, the next time the player chooses a strategy it must be evaluated as strictly worse. Given that the game is finite and that the memory of each player is finite, the evaluation of any strategy cannot keep strictly declining infinitely often.

Hence, if play was to cycle (perhaps, probabilistically), then some player i must return to a state in which she evaluates two (or more) strategies equally. If this were not the case then player i would return to a strategy infinitely often when its evaluation had strictly declined. But this cannot happen by the argument in the preceding paragraph. Hence, if we show that no player can choose a strategy infinitely often because it is evaluated the same as some other strategy, then the proof of the Proposition is complete.

Lemma 2 *No player i can return to a state m in which $f^i\left(m^i_j\right) = f^i\left(m^i_k\right) \geq f^i\left(m^i_l\right)$ for all s_l, $j \neq k \neq l$, an infinite number of times, with probability 1.*

Proof Suppose not. That is, suppose play returns to a state m in which $f^i\left(m^i_j\right) = f^i\left(m^i_k\right) \geq f^i\left(m^i_l\right)$ for some $j \neq k \neq l$, infinitely often. Then, i will choose both s^i_j and s^i_k infinitely often as she chooses each strategy she evaluates the best with positive probability at each time. Let $supp\left(s^i_j\right)$ (resp. $supp\left(s^i_k\right)$) denote the set of payoffs player i remembers from s^i_j (resp. s^i_k) in m and let $\#supp\left(s^i_j\right)$ (resp. $\#supp\left(s^i_k\right)$) denote the number of different payoffs in $supp\left(s^i_j\right)$ (resp. $supp\left(s^i_k\right)$). Clearly, either $\#supp\left(s^i_j\right) > 1$ or $\#supp\left(s^i_k\right) > 1$ or both, as otherwise she could not evaluate s^i_j and s^i_k the same because all her payoffs are distinct. This, however, implies that also some other player(s) must be randomizing among strategies they evaluate the same otherwise player i would not experience different payoffs for a strategy that is used in the cycle. And we know that there is an s^i that is used with $\#supp\left(s^i\right) > 1$. So we have now established that the cycle under consideration has at least two players randomizing. As these randomizations are independent of each other, there is always a positive probability that player i obtains H identical payoffs from H consecutive choices of s^i. From such an "event" the player would

> end up remembering only one payoff from s^i. If the agent continues to evaluate s^i_j and s^i_k the same, consider another sequence of H rounds in which the player obtains the same payoff from the other strategy. Such a sequence of play, which has positive probability, ensures that the player remembers only one payoff from s^i_j and one payoff from s^i_k. Over the infinite repetition of the game, such an event has probability 1 (if the players continue to evaluate the two strategies the same). At that time, or earlier, player i will not evaluate the two strategies the same. Hence, no player can return to a state in which she evaluates two or more strategies as being the best infinitely often. □

Hence, play must at some time settle upon an LME. □

The conclusions reached in Propositions 1 and 2 contrast in a surprising way with the result obtained in Sarin (2000) concerning games against nature. Sarin shows that, in a game against nature, a player converges to choose the strategy that gives her the maxmin payoff. That is, the player converges to her maxmin strategy. First, he shows that the individual cannot converge to play any strategy other than her maxmin strategy. Suppose to the contrary. Then she will experience from this strategy a long enough run of the worst possible payoff from this strategy so that the worst payoff is all the decision maker remembers from this strategy. At this time, or earlier, the individual must evaluate her maxmin strategy (or some other strategy) as being better. Hence, the individual cannot converge to play any strategy other than her maxmin strategy. Next, consider a state in which the player currently evaluates the maxmin strategy as being the best and evaluates every other strategy as being worse than the maxmin strategy could possibly be evaluated. In such a state the individual chooses the maxmin strategy forever. Sarin shows that such a state is reached from any other state with probability one.

Proposition 2 shows that players may converge to play strategies other than their maxmin strategies if the payoffs they obtain by these strategies are higher than their maxmin payoffs. Hence, when facing a game environment agents do (weakly) better than when facing the decision theoretic environment. Intuitively, this happens since players may (unsuspectingly) influence the nonstationary game environment they face. What is surprising is that they only influence it in a way that "improves" it. That is, players only "reinforce" strategy profiles that lead to outcomes better than the maxmin outcomes. Another reason for the superior performance in games is that other players (typically) choose deterministically, whereas in a game against nature, the other player ("nature") chooses stochastically. Hence, players do (weakly) better in nonstationary deterministic environments than in stationary stochastic environments.[14]

4 Stable LME

So far we have studied players whose memory capacity is limited. In this section we shall assume that memory is also imperfect, in the sense that arbitrary items in the memory may be forgotten at any time.[15] Specifically, we shall assume that each item, in each player's memory, is forgotten with some small probability ε, and is forgotten independently of the others.[16] Further, we suppose that if an item associated with strategy s^i is forgotten, then it is replaced by an arbitrary payoff which is obtainable from using this strategy. This can be justified by assuming that the agent has some vague notion of what a strategy can or cannot achieve.[17] We shall refer to the event that one element of the memory is altered as a mistake or a mutation.[18]

Without noise, a player's memory changes only for the strategy she used in the last period, i.e., if she used strategy s^i in period t, then $m(\widetilde{s}^i, t) = m(\widetilde{s}^i, t+1)$ for all $\widetilde{s}^i \neq s^i$. In the presence of mutations, each entry in a player's memory $m^i(t)$ may be altered at each point in time. With such noise in players' memories, we obtain a process P^ε that is aperiodic and irreducible and, therefore, has a unique stationary distribution μ^ε for every $\varepsilon > 0$. We will focus on the limit invariant distribution $\mu^* \equiv \lim_{\varepsilon \to 0} \mu^\varepsilon$. By standard arguments (see, e.g., Kandori, Mailath and Rob (1993), Young (1993)) we know only states which are elements of absorbing sets under P, which in our case are all singletons in generic games, can appear in the support of μ^*. These states are called *stochastically stable states* as only they will be observed with positive probability in the long run. A limited–

[14] It would be interesting to study game environments where each strategy combination is associated with a lottery rather than with deterministic payoffs.

[15] Recall that there are $H \times J_i$ items in player i's memory.

[16] We could assume that items change with different probabilities and, as long as all probabilities are of the same order of magnitude, this would not affect the results. Further alternatives are discussed in the section 6.

[17] This assumption ensures that the perturbed Markov process which we are going to analyse operates on the same state space as the unperturbed process. This helps us keeping the notation simple. The payoff could, for example, be drawn from some distribution over all possible payoffs. Our results would be robust to different replacement rules as long as the payoffs are drawn from a distribution with the same range of payoffs.

[18] There is evidence that items in memory periodically "mutate" (see, e.g., Schacter 1996).

memory equilibrium which is played in a stochastically stable state will be called a *stable limited–memory equilibrium (SLME)*.

Definition 3 *A strategy profile s is an SLME if it is an LME and if there is a state inducing s which has positive probability under μ^*.*

The set of SLME, in general, depends on the exact specification of the evaluation rule f. That is, for a given game Γ and a given size H of players' memories, SLME may be different for different evaluation rules. Furthermore, for a given evaluation rule the set of SLME can be different for games with identical best reply correspondences. To illustrate this, consider a 2x2 game with two strict equilibria $(up, left)$ and $(down, right)$ which are the only LME, i.e. the only strategy profiles played in absorbing states of P. Also, suppose that the evaluation rule is the average rule, that H is larger than one, and that players are locked in equilibrium $(up, left)$.

In order to identify SLME we need to analyze how many mutations are required to switch from one state to the other and how many are required to switch back. Roughly speaking, the state which can be reached with fewer mutations will be the stochastically stable one. So, suppose that one element of the row player's memory changes such that she switches to strategy *down*. The next period's outcome is $(down, left)$. This will change the row player's evaluation of strategy *down* and the column player's evaluation of strategy $left$. However, to what extent the evaluations change depends on the exact payoffs yielded by the strategy combination $(down, left)$. Given that the row player earns less than in the previously played equilibrium she will surely return to play *up*. Now, if the column player's payoff for $(down, left)$ is only slightly less than the $(up, left)$–equilibrium payoff she may continue playing strategy $left$. If, however, the column player's payoff is significantly lower then she may switch to $right$. This can induce further movements, away from the previously played equilibrium. Accordingly, SLME depend not only on the ordinal ranking of payoffs.

Consequently, in this section we restrict our attention to specific evaluation rules. In particular, we consider monotonic evaluation rules which, in contrast to the average rule, induce an ordinal ranking among strategies. In the next section, we study SLME in classes of games without restricting the evaluation rules used by the players.

The following result concerns common interest games in which there is a payoff vector that strongly Pareto dominates all other feasible payoffs. The result shows that if players have a large enough memory (relative to the number of players) and use the optimistic maximum rule then the unique Pareto–efficient outcome is the unique SLME.

Proposition 3 *Suppose that players use the maximum evaluation rule and that $H > n$. Then, in generic n–player common interest games, the strategy profile s inducing the Pareto–optimal outcome is the unique SLME.*

Proof Let Σ be the set of absorbing states in which s is played. We first show that from any state outside of Σ, a state $m \in \Sigma$ can be reached with at most n simultaneous mutations. This happens if each player i replaces one element of $m(s^i, t)$ by $u^i(s)$. This induces the players to switch simultaneously to s^i in the next period as the evaluation of s^i instantaneously assumes its maximum, which is greater than the maximum evaluation of any other strategy. Obviously, once they play s, they will continue to play s such that they will reach a state in Σ with $m(s^i) = U^i(s)$ for all players i. Next, consider how many mutations are necessary to leave Σ. In order to make at least one player i change her strategy, all elements of $m(s^i)$ have to be replaced by values lower than $u^i(s)$, i.e. she has to experience H simultaneous mutations. It follows by standard arguments (see, e.g., Vega-Redondo (1997), Young (1993)) that only states in Σ are stochastically stable. $\square$

This proposition illustrates that players may benefit from being optimistic and having larger memories (or playing with only a few other players). To see the impact of memory size, relative to the number of players, on the result, suppose that $H < n$. The efficient outcome can still be reached with n mutations. However, depending on the exact payoffs, it could now be possible that one player erases her complete memory of payoffs obtained from s^i (which requires less than n mutations), and keeps playing an alternative strategy $\widetilde{s}^i$ for the next H periods. After this time span all other players j will remember only $u^j(\widetilde{s}^i, s^{-i})$ for their equilibrium strategies s^j. As these payoffs may be lower than the maximum of payoffs remembered for some other strategies $\widetilde{s}^j$, it is possible that some of these players also turn away from the equilibrium strategy such that the dynamics will move further away from the efficient outcome. Hence, without knowing more about the payoff function of the game, it is impossible to predict the SLME.

Our next result is also concerned with players using the maximum rule and playing certain games where the interests of the players are aligned. We refer to this class as games with strong common interest.

Definition 4 *A game Γ is a game of strong common interest if it has a unique equilibrium and if $\widetilde{s}^i = \arg\max_{s^i} u^i(s^i, s^{-i})$ implies $u^j(\widetilde{s}^i, s^{-i}) \geq u^j(s^i, s^{-i})$ for all s^i, s^{-i} and j.*

This definition ensures that *all* players' payoffs weakly increase if a single player deviates from a non–Nash strategy profile to her best reply. Hence, individual best replies are in the common interest of all players. It is obvious that in generic games of strong common interest the unique Nash equilibrium is also the unique Pareto–efficient outcome, i.e., the class of games with strong common interest is contained in the class of games with common interest.[19]

Proposition 4 *Suppose that players use the maximum evaluation rule. Then, in generic n–player games of strong common interest, the unique equilibrium s is an SLME. If $H > 1$ then it is also the unique SLME.*

Proof We will show that a state in which the equilibrium is played can be reached from any other state by a sequence of one–shot mutations. The first claim in the Proposition then follows by standard arguments. To construct this sequence, consider any state m in which some LME $s' \neq s$ is played. As s' is not an equilibrium there exists at least one player i who could obtain a higher payoff by deviating to her best reply. One mutation is sufficient to induce this deviation. As soon as this occurs, she will switch to her best reply while all other players will continue to play s'^{-i}. (Their new payoffs have increased). Thus, the dynamics will reach a new LME. An arbitrary amount of time can pass. And, if the new LME is also not a the unique Nash equilibrium, a further single mutation can be constructed in the same way. Eventually, by finiteness, the dynamics will reach an absorbing state inducing s. In order to prove the second claim, suppose players are currently in a state inducing s. It is sufficient to show that s cannot be left with a sequence of one-shot mutations. Imagine a mutation which makes one player switch her strategy. Obviously, this will decrease her payoff, and, as she still remembers the equilibrium payoff, she will immediately switch back. Furthermore, all other players will experience payoffs lower than their equilibrium payoffs for their equilibrium strategies. However, as $H > 1$, they still remember at least one equilibrium payoff. As they evaluate their memory by the maximum rule they will return to the equilibrium strategy. Hence, after a single mutation the dynamics always lead back to a state inducing s.$\square$

The last part of the proof illustrates the role of memory size. If $H > 1$, a single deviation cannot make the equilibrium payoffs forgotten. With $H = 1$,

[19] Our definition of games of strong common interest is related to Monderer and Shapley's (1996) definition of potential games.

a single mutation may make player i deviate from her equilibrium strategy and induce others to also move away from the equilibrium strategy (in the next round).

Both of our results on SLME depend on players being optimistic. Our next result reveals that when players are pessimistic elements of risk avoidance may have a strong enough impact to prevent them from coordinating on a Pareto–dominant equilibrium. The next result considers players who are pessimistic and use the minimum evaluation rule. It shows that being pessimistic may lead to less efficient outcomes even if memories are large. Note that the result applies for all generic games and all memory sizes. Let $s_{\max\min}$ denote the strategy profile in which each player plays her maxmin strategy.

Proposition 5 *Suppose players use the minimum rule. Then, in generic games, $s_{\max\min}$ is an SLME. If $s_{\max\min}$ is, in addition, a Nash equilibrium then it is the unique SLME.*

Proof In order to prove the first statement, we show that a state in which all players play their maxmin strategies can be reached from any other state by a sequence of one–shot mutations. The claim then follows by standard arguments. Consider a sequence of one–shot mutations $k = 1, 2, ..., \sum_i (J_i - 1)$. Each mutation k replaces one element of player i's memory of her strategy $s^i_j \neq s^i_{\max\min}$. Specifically, each mutation replaces one item (or, payoff) of player i's memory of strategy j by the minimum payoff that strategy can give. Once the sequence has been completed, each player evaluates each strategy according to the minimum payoff it can give. This leads each player to choose her maxmin strategy.

In order to prove the second statement, we show that the state in which all players play their maxmin strategies requires at least two simultaneous mutations to be left. Suppose a single mutation would be sufficient to make player i switch from her maxmin strategy to a different s^i_j. Note that this can only occur if player i replaces an item in her memory for s^i_j, i.e., if her evaluation of s^i_j suddenly improves.[20] After this switch, all other players will continue to play their maxmin strategies as they get a payoff not smaller than their maxmin payoff and as they remember at least one smaller payoff from each of their other strategies. Since mutual maxmin is, by assumption, a Nash equilibrium it

[20] The evaluation of the maxmin strategy cannot fall below the maxmin payoff.

is a strict equilibrium, because the game is generic. Consider player i: s_j^i will give her a strictly lower payoff than her maxmin strategy has given her previously (which she still remembers). Hence, she will switch back to her maxmin strategy. Thus, the dynamics lead always back into mutual maxmin play if there is only a single mutation.$\square$

Whereas Propositions 3 and 4 suggest that optimism (as implied by the use of the maximum evaluation rule) leads to efficiency in some games,[21] Proposition 5 shows that pessimism (as implied by the use of the minimum rule) leads to potentially very bad outcomes in all games. Proposition 4 holds for all memory capacities. This implies that it also holds for the limiting case of $H = 1$, when all evaluation rules collapse into one. We summarize two consequences of Proposition 3 for two much studied games in the following corollary:[22]

Corollary 1 *Suppose players use the minimum rule or that $H = 1$. Then the following statements hold:*

a) In 2x2 Prisoners' Dilemma games, the unique SLME is given by mutual defection.

b) In symmetric 2x2 coordination games in which payoff dominance and risk dominance do not select the same equilibrium, the risk dominant equilibrium is the unique SLME.

Proof a) As mutual defection is an equilibrium in dominant strategies the claim follows immediately. b) Consider the game below with $d > b$, $a > c$, and $d > a$.

	$left$	$right$
up	a, a	b, c
$down$	c, b	d, d

$(down, right)$ is the payoff dominant equilibrium and $(up, left)$ is the risk dominant equilibrium if $a + b > c + d$, in which case $b > c$. Hence, up and $left$ are the maxmin strategies. $\square$

5 Iterated Uniform Dominance

We begin this section with some definitions.

[21] Nice properties of optimism have recently been shown in a number of papers including Gilboa and Schmeidler (1996) and Sarin and Vahid (1999).

[22] Note that the first statement can easily be extended to n–person PD games where each player has a dominant strategy.

Definition 5 *A strategy* s^i_j *is* uniformly dominated *if there exists another strategy* s^i_k *such that* $u^i_{\min}\left(s^i_k\right) > u^i_{\max}(s^i_j)$.

That is, we say that strategy s^i_j is uniformly dominated by strategy s^i_k if the minimum payoff the latter can give is greater than the maximum payoff s^i_j can give. Hence, while cooperation is (strictly) dominated by defection in the Prisoner's Dilemma, it is not uniformly dominated. Consider the game below in which the row player has three strategies (u, m, d) and the column player has two strategies (l, r).

	l	r
u	4,4	6,2
m	2,1	3,6
d	5,3	2,2

In this game, m is uniformly dominated for the row player. No other strategies are uniformly dominated for either player.

We now develop the definition of the set of strategies that survive the iterated elimination of uniformly dominated (IUD) strategies. Let $\tilde{S}^{i,1}$ be obtained from S^i by deleting from the latter all strategies that are uniformly dominated. Let $\tilde{S}^1 = \times_{i\in I}\tilde{S}^{i,1}$ denote the set of strategy profiles that may be played after each player has removed the uniformly dominated strategies. It is natural to call $\tilde{S}^1$ the set of strategy profiles that survive one round of removal of uniformly dominated strategies. Clearly, this set is non–empty. In particular, in the Prisoner's Dilemma no strategy is eliminated by one round of removal of uniformly dominated strategies. In the game above, one round of removal of strategies that are uniformly dominated leads to the following reduced game

	l	r
u	4,4	6,2
d	5,3	2,2

Next, we construct the set of strategies that survive the elimination of uniformly dominated strategies in $\tilde{S}^1$ and call this set of strategies $\tilde{S}^2$. In the above game this is given by

	l
u	4,4
d	5,3

We similarly construct $\tilde{S}^3$, $\tilde{S}^4$, ... Observe that $\tilde{S}^{n+1} \subset \tilde{S}^n$ for any n.

Definition 6 *The set of strategies which survive the iterated elimination of uniformly dominated strategies (SIUD) is* $\tilde{S}^\infty \equiv \cap_{n=1}^\infty \tilde{S}^n$.

Clearly, $\tilde{S}^\infty$ is non–empty. In particular, it is obviously at least as large as the set of strategies which survive the iterated removal of (strictly) dominated strategies (SISD), which is known to be non–empty. It is equal to S in games in which there does not exist a uniformly dominated strategy, as is the case in the Prisoner's Dilemma. In the above game, however, the iterated elimination of uniformly dominated strategies results in the unique strategy profile $(d,\, l)$, which is also the unique Nash equilibrium.

Definition 7 *A game is solvable by the iterated elimination of uniformly dominated strategies if* $\tilde{S}^\infty$ *is a singleton.*

It is easily seen that if a game is solvable by the iterated elimination of uniformly dominated strategies, then the game has a unique Nash equilibrium. This is because the SISD is contained in the SIUD. As the former is well known to be non–empty in all games, we know that if the latter is a singleton, then the former must be also. But, we also know that Nash equilibrium coincides with the SISD if it is a singleton. Hence, when the SIUD is a singleton, it coincides with the Nash equilibrium.

Next, we provide a result providing some insight into how players with limited memories play uniform dominance solvable games.

Proposition 6 *In uniform dominance solvable games the unique equilibrium s is an SLME.*

Proof We will show that a state in which the equilibrium s is played can be reached from any other state with a series of one–shot mutations. From this the claim follows immediately. Let $I(l)$ be the set of players which eliminates strategies in the lth iteration of eliminating uniformly dominated strategies and let $J(l)$ be the set of eliminated strategies. Note first that strategies in $J(1)$ are never played. Nevertheless, players can remember outcomes in which strategies in $J(1)$ are played, for example, because of the initial conditions. We start the construction of the sequence of one–shot mutations by "erasing" these memories. More precisely, we replace all $m_h(s^i_j, t)$ which resulted from the use of strategies in $J(1)$ by a payoff that strategy s^i_j can obtain in $\widetilde{S}^1$. Note that this need not happen simultaneously. Rather, an arbitrary amount of time can pass between mutations. When all payoffs stemming from strategies in $J(1)$ have been replaced, it is obvious that players in $I(2)$ will no longer play strategies in $J(2)$. (Strategies in $J(2)$ have not been eliminated in the first round because they give a maximal payoff higher than the minimum payoff of all other strategies. However, as they can be eliminated in the second round, it is clear that they give this maximal payoff only against strategies in $I(1)$. Otherwise they could not be eliminated in the second round. But due to the mutations players in $I(2)$ have now "forgotten" those payoffs and remember only payoffs whose maximum is lower than the minimum of the payoffs they remember from some other strategy. Hence, they do not play strategies in $J(2)$.) In the next subsequence of one–shot mutations all players' memories of payoffs resulting from strategy combinations containing strategies in $J(2)$ are replaced in the same fashion. As a consequence players in $I(3)$ will no longer use strategies in $J(3)$. The sequence of mutations is completed by repeating the same steps until, eventually, players play equilibrium s.$\square$

The above result places restrictions neither on the memory size of players nor on the evaluation rule they use. However, without further assumptions, we cannot show that players will play the equilibrium all the time as we cannot prove uniqueness for the above case. The reason for this can be easily illustrated. Suppose player i experiences a single mutation which makes her switch from s^i to some other strategy $\widetilde{s}^i$. As $\widetilde{s}^i$ is not a best response against s^{-i}, it may happen that she immediately returns to s^i. However, the single instance of her playing $\widetilde{s}^i$ may have caused other players to re–evaluate their equilibrium strategies. As a consequence of this, several players other than i may deviate from the equilibrium strategy in the following period (even if i herself has returned). Thus, to prove uniqueness, we need to know more specific details of the game.

Next, we consider the class of games which are solvable by (standard) iterative elimination of dominated strategies. As this class contains the class of uniform dominance solvable games, it is not surprising that an analogous result requires additional assumptions. The next proposition shows that the Nash equilibrium in such games is an SLME if players use the minimum evaluation rule.

Proposition 7 *In dominance solvable games the unique equilibrium s is an SLME if players use the minimum rule (regardless of memory size) or if $H = 1$.*

Proof The proof is analogous to the one of Proposition 6, and we use the same notation. The main difference between dominance solvable games and games solvable by uniform dominance is that in the former players in $I(1)$ may use strategies in $J(1)$. More generally, players in $I(k)$ may use strategies in $J(k)$ in the reduced game $\widetilde{S}^{k-1}$. Thus, the sequence of one–shot mutations to reach s has to be more elaborated. We start the sequence of mutations by replacing for each strategy $s^i_j \in J(1)$ one of the remembered payoffs by $u^i_{\min}(s^i_j)$. Between two mutations an arbitrary amount of time can pass. When all mutations of this first subsequence have occurred, players $i \in I(1)$ who played strategies in $J(1)$ will have switched to strategies in $\widetilde{S}^{i,1}$ (where $\widetilde{S}^{i,1}$ is now defined by applying the standard notion of strict dominance). Next we proceed—as in the proof of Proposition 6—by replacing all $m_h(s^i_j, t)$ which resulted from the use of a strategy in $J(1)$ by a payoff obtainable in $\widetilde{S}^1$. However, this does not yet ensure that players $i \in I(2)$ do not use strategies in $J(2)$. In order to make them switch to strategies in $\widetilde{S}^{i,2}$ additional mutations are required. But, as in the first subsequence, one mutation per player is sufficient to make them switch. If one remembered payoff of a strategy $s^i_j \in J(2)$ is replaced by $u^i_{\min}(s^i_j)$ there must be a strategy $s^i_k \in \widetilde{S}^{i,2}$ for which a greater payoff is remembered. Due to the minimum evaluation rule this implies that the player switches to this strategy. As soon as the new strategy profile is in $\widetilde{S}^2$ we again proceed as in the proof of Proposition 6. This procedure can be repeated until the equilibrium is reached.□

6 Discussion

There are two idealizations about memory that form the core of our model. First, there is the particular model of limited memory capacity. According to it, only a finite number of items could be remembered, and agents remembered the same number of items from each strategy, namely the most recently experienced payoffs from the respective strategies. Second, there is the specific manner in which this finite memory capacity is imperfect: Each item in the memory of a player is forgotten with small positive probability which is independent across items, and the forgotten items are replaced by other plausible items. We turn now to discuss these idealizations about memory.

Given our finite brains and neural content there is compelling case to suppose that memory capacity is limited. The assumption that the decision maker remembers the same number of payoffs from each strategy was made largely for convenience—it made comparing (or evaluating) alternative strategies particularly easy since it only involved comparing vectors of payoffs of the same dimension. In certain contexts, however, it seems plausible that a player remembers a different number of payoffs from each of her strategies. Also, different players probably remember different amounts of information. Therefore, it would be nice if our results were to extend to the situation in which the number of payoffs an agent has in her memory depended on who she was and on the strategy. That is, if H was replaced by a player and strategy specific, H_j^i. We have not pursued such an extension in this paper.

In some situations, it might be reasonable to assume that the amount of memory allocated to strategies that are not played for a long while begins to decay. This extension is allowed for in Sarin (2000), in which he argues that the maxmin result is robust to assuming such decay. Further analysis is required to see how this change would affect the current model.

The assumption that the payoffs remembered are the most recently experienced payoffs from those strategies was to reflect the intuition that "more recent happenings are better remembered" without giving up the first assumption. Now suppose that an agent remembers only the most recent K payoffs, irrespective of the strategy that was chosen. This requires us to address how agents evaluate strategies from which they recall no payoffs. Potential ways of doing this are discussed in Sarin (2000).

The second important aspect of the analysis of this paper is that we suppose that the items that are stored in memory are imperfectly stored. There is plenty of evidence to suggest that people possibly forget any item in their memory (see, e.g., Schacter 1996). More specifically, our model supposed that each item in memory was forgotten with a small fixed probability. Intuitive arguments might suggest that agents forget more recent items less frequently than those stored earlier[23] and we could have supposed that items stored earlier are forgotten with higher probability, as long as all probabilities are of the same order of magnitude. While we believe that our assumption that items are forgotten independently of one another and independently of the current state is a useful first approximation, there might be reasons to make different assumptions for specific applications. In that case one might want to introduce "state–dependent" forgetfulness.[24]

Another assumption we made was that an item that was forgotten was

[23] Of course, this is partially reflected in how agents utilize their memory in the basic model. Recent payoffs are remembered, earlier payoffs are forgotten.

[24] This may change the standard results (see, e.g., Bergin and Lipman 1996).

replaced, and furthermore, that it was replaced by an item that could possibly have been forgotten. We assumed that forgotten items are replaced in this manner to ensure that the number of remembered payoffs remained constant for all strategies. We assumed realistic replacement as it allowed us to work with the same finite state space and we thought it to be an useful starting point in the study of how memory might affect behavior.

7 Conclusion

The present study takes a first step in modelling how players with memory limitations, who have very little information about their choice environment, play games. Our preliminary results suggest, firstly, that players with limited memories tend to be cautious. This cautiousness is reflected by the emergence of maxmin strategies in various settings. Second, players with memory limitations tend to do weakly better in games with other players than in a game with nature. Whereas in decision problems the player converges to their maxmin payoffs, in games players converge to at least their maxmin payoffs. Third, players with severe memory limitations ($H = 1$) achieve quite a lot. For example, they learn to use dominant strategies and will play equilibria in dominance solvable games. However, they fail to coordinate on Pareto–efficient equilibria in common interest games and end up playing the risk–dominant equilibrium. Pessimistic players who use the minimum rule achieve very similar things as players with minimal memory. For more optimistic players we have seen that a comparatively large memory may improve their performance.

Finally, we have seen that in classes of games with especially obvious solutions, i.e., in games which are solvable by iterated uniform dominance, the behavior of players with limited memories is robust to memory sizes and evaluation rules. For games with less obvious solutions this is not true.

References

1. Aumann, R. and S. Sorin (1989): "Cooperation and bounded recall," *Games and Economic Behavior*, 1, 5-39.

2. Barbera, S. and M. Jackson (1988): "Maximin, Leximin, and the protective criterion: Characterizations and Comparisons," *Journal of Economic Theory*, 46, 34-44.

3. Bergin, J., and B. Lipman (1996): "Evolution with state–dependent mutations," *Econometrica*, 64, 943-956.

4. Chen, Y. (1999): "Asynchrony and learning in serial and average cost pricing mechanisms: An experimental study," mimeo, University of Michigan.

5. Conlisk, J. (1996): "Why bounded rationality," *Journal of Economic Literature*, 34, 669-700.

6. Dow, J. (1991): "Search decisions with limited memory," *Review of Economic Studies*, 58, 1-14.

7. Ellison, G. (1997): "Learning from personal experience: One rational guy and the justification of myopia," *Games and Economic Behavior*, 19, 180-210.

8. Friedman, E. and S. Shenker (1998): "Learning and implementation on the internet," mimeo, Rutgers University.

9. Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141-153.

10. Gilboa, I. and D. Schmeidler (1996): "Case–based optimization," *Games and Economic Behavior*, 15, 1-26.

11. Greenwald, A., E. Friedman and S. Shenker (1998): "Learning in network contexts: Experimental results from simulations," *Games and Economic Behavior*, forthcoming.

12. Hurkens, S. (1995): "Learning by forgetful players," *Games and Economic Behavior*, 11, 304-329.

13. Kandori, M., G. Mailath, R. Rob (1993): "Learning, mutation, and long run equilibria in games," *Econometrica*, 61, 29-56.

14. Lehrer, E. (1988): "Repeated games with stationary bounded recall strategies," *Journal of Economic Theory*, 46, 130-144.

15. Lehrer, E. (1994): "Finitely many players with bounded recall in infinitely repeated games," *Games and Economic Behavior*, 7, 390-405.

16. Mullainathan, S. (1998): "A memory based model of bounded rationality," mimeo, Harvard University.

17. Osborne, M., and A. Rubinstein (1998): "Games with procedurally rational players," *American Economic Review*, 88, 834-847.

18. Sabourian, H. (1998): "Repeated games with M-period bounded memory (pure strategies)," *Journal of Mathematical Economics*, 30, 1-35.

19. Sarin, R. (2000): "Decision rules with bounded memory," *Journal of Economic Theory*, 90, 151-160.

20. Sarin, R., and F. Vahid (1999): "Payoff assessments without probabilities: A simple dynamic model of choice," *Games and Economic Behavior*, 28, 294-309.

21. Schacter, D. (1996): *Searching for Memory: The Brain, the Mind and the Past*, Basic Books.

22. Sela, A. and D. Herreiner (1999): "Fictitious play in coordination games," *International Journal of Game Theory*, 28, 189-198.

23. Sethi, R. (2000): "Stability of equilibria in games with procedurally rational players," *Games and Economic Behavior*, 32, 85-104.

24. Sonsino, D. (1998): "On subjective uncertainty and myopic behavior," mimeo, Technion.

25. Vega-Redondo, F. (1997): "The evolution of Walrasian behavior," *Econometrica*, 65, 375-384.

26. Young, P. (1993): "The evolution of conventions," *Econometrica*, 61, 57-84.